

BAB 1 PENDAHULUAN

1.1 Latar Belakang Masalah

Secara epidemiologis, gangguan tidur menunjukkan angka kejadian yang cukup tinggi baik pada tingkat nasional maupun global. Survei pada populasi umum di Indonesia melaporkan bahwa sekitar 28 juta penduduk atau kurang lebih 10% dari total populasi mengalami gangguan tidur, yang mencerminkan bahwa permasalahan ini telah menjadi isu kesehatan masyarakat yang signifikan [1]. Pada tingkat global, prevalensi insomnia diperkirakan berkisar antara 10–30% pada populasi dewasa dunia, dengan estimasi terbaru menunjukkan angka sebesar 16,2% atau setara dengan lebih dari 850 juta orang; sekitar 7,9% di antaranya termasuk dalam kategori insomnia berat yang berdampak nyata terhadap fungsi sosial dan produktivitas sehari-hari [2]. Selain insomnia, gangguan tidur lain seperti *Obstructive Sleep Apnea* (OSA) juga memiliki prevalensi yang tinggi. Data klinis menunjukkan bahwa OSA dapat ditemukan pada sekitar 9–24% pria dewasa dan 4–9% wanita dewasa usia 30–60 tahun, dengan angka yang lebih tinggi pada kelompok usia lanjut serta individu dengan obesitas [3].

Gangguan tidur atau *sleep disorder* merupakan kondisi medis yang ditandai oleh perubahan dalam kualitas, durasi, dan pola tidur yang tidak sesuai dengan kebutuhan fisiologis tubuh, yang pada akhirnya berdampak negatif pada fungsi kesehatan secara menyeluruh selama periode siang hari. Secara normal, tidur memainkan peran penting dalam pemulihan fisik dan mental, termasuk pengaturan sistem imun, fungsi kognitif, serta stabilitas emosional [4]. Salah satu jenis gangguan tidur adalah insomnia, yaitu kesulitan dalam memulai atau mempertahankan tidur yang baik sehingga mengakibatkan berbagai gangguan fungsi pada saat siang hari [2], sedangkan *Obstructive Sleep Apnea* (OSA) merupakan kondisi terhambatnya saluran napas terhambat berulang selama tidur sehingga menyebabkan sesak napas, hipoksia intermiten, dan fragmentasi tidur yang mengakibatkan kelelahan siang hari serta penurunan kualitas hidup [5]. Gangguan tidur ini bukan sekadar gejala sementara, melainkan memiliki konsekuensi luas bagi kesehatan; kondisi ini dapat memicu penyakit kardiovaskular seperti penyakit jantung dan stroke [6], kanker, serta gangguan kesehatan mental seperti depresi, gangguan kecemasan, dan bipolar [1]. Selain itu, gangguan tidur

juga diidentifikasi sebagai faktor predisposisi terhadap dysregulasi metabolik dan tekanan darah tinggi yang berkontribusi pada hipertensi dan diabetes mellitus tipe 2 [7]. Hubungan ini bersifat timbal balik karena penyakit kardiovaskular yang sudah ada pun dapat memperparah disfungsi tidur sehingga memperbesar beban penyakit secara keseluruhan [8]. Di luar konsekuensi kardiometabolik, gangguan tidur kronis juga mempercepat proses neurodegeneratif melalui mekanisme stres oksidatif dan peradangan sistemik yang meningkatkan risiko demensia dan Alzheimer [9]. Dampak-dampak ini menunjukkan bahwa gangguan tidur bukan sekadar gangguan minor; jika tidak ditangani sejak dini melalui deteksi yang komprehensif, dampaknya dapat meluas ke berbagai aspek kesehatan dan menimbulkan beban klinis serta sosial yang signifikan.

Meskipun berbagai penelitian telah menunjukkan bahwa machine learning dapat meningkatkan akurasi identifikasi gangguan tidur dibandingkan metode konvensional, sebagian besar studi masih menerapkan model secara individual tanpa evaluasi komparatif yang terstruktur. Penelitian berbasis Random Forest menunjukkan hasil klasifikasi yang baik pada dataset *Sleep Health and Lifestyle*, namun tidak membandingkannya secara langsung dengan algoritma *ensemble* lain dalam satu skenario eksperimen yang sama [10]. Pendekatan XGBoost juga telah digunakan dan terbukti mampu mengatasi keterbatasan metode tradisional dalam identifikasi gangguan tidur seperti sleep apnea melalui data rekam medis elektronik [11], sementara studi lain mengevaluasi berbagai model regresi dan klasifikasi untuk memprediksi gangguan tidur yang berkaitan dengan kondisi fisik dan emosional, dengan hasil yang bervariasi antar algoritma [12]. Beberapa penelitian juga turut menguji berbagai algoritma untuk deteksi insomnia dan gangguan tidur [13, 14, 15], namun umumnya hanya membandingkan performa model secara terpisah tanpa mengeksplorasi potensi penggabungan model.

Terdapat perbedaan mendasar antara dua paradigma *ensemble* yang paling banyak digunakan, yaitu *bagging* yang direpresentasikan oleh Random Forest dan *boosting* yang direpresentasikan oleh XGBoost dengan masing-masing memiliki mekanisme pembangunan model dan penanganan *bias-variance* yang berbeda [10, 11]. Di sisi lain, pendekatan *stacking ensemble* telah terbukti mampu melampaui performa model tunggal maupun metode *ensemble* konvensional lainnya dengan cara memanfaatkan kekuatan prediktif dari beberapa *base learner* secara bersamaan. Hal ini telah didemonstrasikan pada domain yang berkaitan dengan tidur, dengan bukti bahwa *stacking* yang menggabungkan XGBoost sebagai *base learner* terpilih berhasil mencapai akurasi 94,48% dan melampaui seluruh

model *ensemble* pembandingan dalam tugas klasifikasi posisi tidur [16]. Meskipun domain tersebut berbeda dari klasifikasi gangguan tidur, temuan ini menunjukkan potensi *stacking* dalam memanfaatkan keunggulan komplementer antara paradigma *bagging* dan *boosting* pada permasalahan klasifikasi multiclass yang berkaitan dengan tidur. Lebih jauh, efektivitas *stacking ensemble* juga telah dikonfirmasi lintas domain, karena model *stacking* terbukti secara konsisten meningkatkan kemampuan klasifikasi dengan mengintegrasikan keunggulan beberapa model yang beragam [17]. Oleh karena itu, masih terdapat celah penelitian berupa belum adanya komparasi sistematis antara Random Forest dan XGBoost sebagai dua paradigma *ensemble* yang berbeda dalam satu kerangka eksperimen yang konsisten pada domain klasifikasi gangguan tidur, sekaligus belum adanya pengujian pendekatan *stacking ensemble* yang secara khusus menggabungkan kedua paradigma tersebut untuk mengeksplorasi potensi peningkatan performa klasifikasi gangguan tidur secara *multiclass*.

1.2 Rumusan Masalah

Berdasarkan latar belakang dan kebutuhan penelitian, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana implementasi dan evaluasi algoritma *Random Forest* dan XGBoost sebagai dua paradigma *ensemble* yang berbeda dalam mengklasifikasikan gangguan tidur?
2. Apakah penerapan metode *stacking ensemble* yang menggabungkan *Random Forest* dan XGBoost sebagai *base learner* dapat meningkatkan performa klasifikasi dibandingkan dengan penggunaan model tunggal?

1.3 Batasan Permasalahan

Agar penelitian ini lebih terarah dan tidak meluas, maka batasan masalah yang ditetapkan adalah sebagai berikut:

1. Penelitian ini hanya membandingkan dua algoritma *supervised learning* berbasis *ensemble*, yaitu *Random Forest* (paradigma *bagging*) dan XGBoost (paradigma *boosting*).
2. Selain model tunggal, penelitian ini hanya menggunakan satu teknik *ensemble*, yaitu *stacking*.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan, maka tujuan dari penelitian ini adalah sebagai berikut:

1. Menganalisis dan membandingkan kinerja algoritma *Random Forest* dan XGBoost sebagai representasi dari dua paradigma *ensemble* yang berbeda, yaitu *bagging* dan *boosting*, dalam mengklasifikasikan dataset gangguan tidur.
2. Menerapkan metode *stacking ensemble* dengan *Random Forest* dan XGBoost sebagai *base learner* untuk mengevaluasi potensi peningkatan performa yang dihasilkan dari penggabungan kedua model tersebut.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Memberikan kontribusi dalam pengembangan kajian machine learning, khususnya pada penerapan metode klasifikasi untuk prediksi gangguan tidur.
2. Menjadi referensi bagi penelitian selanjutnya yang membahas komparasi algoritma dan penerapan teknik ensemble learning.
3. Mendukung pengembangan sistem pendukung keputusan dalam bidang kesehatan.

1.6 Sistematika Penulisan

Sistematika penulisan laporan penelitian ini disusun dalam lima bab dengan uraian sebagai berikut.

- Bab 1 PENDAHULUAN membahas latar belakang masalah yang mendasari penelitian, yaitu tingginya prevalensi gangguan tidur dan keterbatasan penelitian terdahulu dalam mengkomparasikan algoritma *ensemble* secara sistematis. Bab ini juga memuat rumusan masalah, batasan permasalahan, tujuan penelitian, manfaat penelitian, serta sistematika penulisan.

- Bab 2 LANDASAN TEORI

menguraikan teori-teori yang menjadi fondasi penelitian, meliputi gangguan tidur (*sleep disorder*), *machine learning*, *ensemble learning*, *Random Forest*, *Extreme Gradient Boosting* (XGBoost), *stacking ensemble*, evaluasi model klasifikasi, serta tinjauan terhadap penelitian-penelitian terdahulu yang relevan.

- Bab 3 METODOLOGI PENELITIAN

menjelaskan alur metodologi penelitian yang terdiri dari delapan tahapan, yaitu studi literatur, pengumpulan data, *exploratory data analysis* (EDA), penggabungan dataset, *preprocessing*, seleksi fitur, pemodelan, dan evaluasi model.

- Bab 4 HASIL DAN DISKUSI

menyajikan hasil eksperimen dari ketiga model yang dibangun, yaitu *Random Forest*, XGBoost, dan *stacking ensemble*, beserta analisis perbandingan performa antar model berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, serta pembahasan terhadap temuan penelitian dibandingkan dengan penelitian terdahulu.

- Bab 5 SIMPULAN DAN SARAN

berisi simpulan yang diperoleh dari hasil penelitian serta saran untuk pengembangan penelitian selanjutnya.

